GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 24, 2002, 22:08:57 ; Search time 68 Seconds
(without alignments)
811.261 Million cell updates/sec

Title: US-09-708-724a-2
Perfect score: 2187
Sequence: 1 MGPWSVVVLVCGMRGLGQAL.....LLAVTREGLERRIISRKRAE 414

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	259.5	11.9	334	22	ABG04612	Novel human diagno
2	256	11.7	133	22	ABB41145	Peptide #8651 enco
3	256	11.7	133	22	AAM62001	Human brain expres
4	256	11.7	133	22	AAM74805	Human bone marrow
5	256	11.7	133	22	AAM34921	Peptide #8958 enco
6	256	11.7	133	23	ABG44596	Human peptide enco
7	254	11.6	122	22	ABG04611	Novel human diagno
8	130	5.9	1178	22	ABB68342	Drosophila melanog
9	129.5	5.9	641	22	AAM38978	Human polypeptide
10	129.5	5.9	698	22	AAM38977	Human polypeptide
11	129.5	5.9	698	22	AAM39032	Human polypeptide
12	129.5	5.9	705	22	AAB94135	Human protein sequ
13	129.5	5.9	706	22	AAM38979	Human polypeptide
14	129.5	5.9	1837	21	AAY85564	Human homologue of
15	128.5	5.9	828	22	ABB65774	Drosophila melanog
16	128	5.9	1593	23	AAM48935	Murine MEKK1-2. M
17	125.5	5.7	595	22	AAE01114	Human gene 1 encod
18	125.5	5.7	595	23	ABG64591	Human albumin fusi
19	125.5	5.7	1560	21	AAB18792	The human ribosome
20	125	5.7	1328	22	AAM78519	Human protein SEQ
21	125	5.7	1331	22	AAM79503	Human protein SEQ
22	123.5	5.6	2639	22	ABG15016	Novel human diagno
23	120.5	5.5	608	22	AAB20164	Human protein asso
24	118	5.4	482	22	AAB31195	Amino acid sequenc
25	118	5.4	482	22	AAY97586	Human secreted pro
26	118	5.4	482	22	AAB65298	Human PRO7170 prot
27	118	5.4	482	22	AAB27225	Human EXMAD-3 SEQ
28	118	5.4	538	22	AAM38985	Human polypeptide
29	118	5.4	538	22	AAE06598	Human protein havi
30	117	5.3	905	18	AAW31186	Human p160 polypep
31	116	5.3	561	22	AAM40771	Human polypeptide
32	115	5.3	503	22	AAB20165	Human protein asso
33	114	5.2	2161	22	AAM78959	Human protein SEQ
34	114	5.2	2189	22	AAM79943	Human protein SEQ
35	114	5.2	2523	22	AAU03503	Human protein kina
36	113.5	5.2	401	21	AAB41664	Human ORFX ORF1428
37	113.5	5.2	425	22	AAU55309	Propionibacterium
38	113.5	5.2	1721	19	AAW48299	Cryptosporidium pa
39	113	5.2	1145	22	AAU04895	Micromonospora eve
40	113	5.2	2429	23	AAE21713	Human PKIN-8 prote
41	112.5	5.1	584	23	AAE23791	Human BCAS1 (breas
42	112.5	5.1	594	22	ABG16928	Novel human diagno
43	112.5	5.1	914	22	ABB69998	Drosophila melanog
44	112.5	5.1	1331	22	ABG28241	Novel human diagno
45	112	5.1	712	23	ABP61500	Human NF-kB activa

ALIGNMENTS

RESULT 1
ABG04612
ID ABG04612 standard; Protein; 334 AA.
XX
AC ABG04612;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #4603.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS68799.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

```
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 34971; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   334 AA;

  Query Match             11.9%;  Score 259.5;  DB 22;  Length 334;
  Best Local Similarity   43.4%;  Pred. No. 4.8e-13;
  Matches   63;  Conservative   12;  Mismatches   29;  Indels   41;  Gaps    4;

Qy     236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
           :||: |||| ||||   || ||||||:|| : ||||||| || |||  ||||:||| |:|
Db     104 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 163

Qy     296 TIAGEPLGHCTFTIS--------PAFVH-------SVLNKRKRQ---------------- 324
            ||||| | | | ||          |:|        :||: |
Db     164 AIAGEPSGPCAFPISVRPDSNNPQQFIHEHTPLEFKLLNELKTSVVNIGVQSPFTLGLPE 223

Qy     325 -----LELLLREVEWPGRGHMAATC 344
                : ||  :|:     | | ||
Db     224 SAFGAMRLLPFDVK-----HWARTC 243


RESULT 2
ABB41145
ID   ABB41145 standard; Peptide; 133 AA.
XX
AC   ABB41145;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #8651 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 27; SEQ ID NO 33780; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a peptide encoded by a single exon
CC   nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   133 AA;

  Query Match             11.7%;  Score 256;  DB 22;  Length 133;
  Best Local Similarity   68.0%;  Pred. No. 2.7e-13;
  Matches   51;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
           :||: |||| ||||   || ||||||:|| : ||||||| || |||  ||||:||| |:|
Db      54 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 113

Qy     296 TIAGEPLGHCTFTIS 310
            ||||| | | | ||
Db     114 AIAGEPSGPCAFPIS 128


RESULT 3
AAM62001
ID   AAM62001 standard; Protein; 133 AA.
XX
AC   AAM62001;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe encoded protein SEQ ID NO: 34106.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
```

PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 34106; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 133 AA;

Query Match 11.7%; Score 256; DB 22; Length 133;
Best Local Similarity 68.0%; Pred. No. 2.7e-13;
Matches 51; Conservative 6; Mismatches 18; Indels 0; Gaps 0;

Qy 236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
:||: |||| |||| || ||||||:|| : ||||||| || ||| ||||:||| |:|
Db 54 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 113

Qy 296 TIAGEPLGHCTFTIS 310
||||| | | | ||
Db 114 AIAGEPSGPCAFPIS 128

RESULT 4
AAM74805
ID AAM74805 standard; Protein; 133 AA.
XX
AC AAM74805;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 35111.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 35111; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 133 AA;

Query Match 11.7%; Score 256; DB 22; Length 133;
Best Local Similarity 68.0%; Pred. No. 2.7e-13;
Matches 51; Conservative 6; Mismatches 18; Indels 0; Gaps 0;

Qy 236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
:||: |||| |||| || ||||||:|| : ||||||| || ||| ||||:||| |:|
Db 54 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 113

Qy 296 TIAGEPLGHCTFTIS 310
||||| | | | ||
Db 114 AIAGEPSGPCAFPIS 128

RESULT 5
AAM34921
ID AAM34921 standard; Protein; 133 AA.
XX
AC AAM34921;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #8958 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 35190; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 133 AA;

Query Match 11.7%; Score 256; DB 22; Length 133;
Best Local Similarity 68.0%; Pred. No. 2.7e-13;
Matches 51; Conservative 6; Mismatches 18; Indels 0; Gaps 0;

Qy 236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
:||: |||| |||| || ||||||:|| : ||||||| || ||| ||||:||| |:|
Db 54 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 113

```
Qy    296 TIAGEPLGHCTFTIS 310
          ||||| | | | ||
Db    114 AIAGEPSGPCAFPIS 128


RESULT 6
ABG44596
ID   ABG44596 standard; Peptide; 133 AA.
XX
AC   ABG44596;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human peptide encoded by genome-derived single exon probe SEQ ID 34261.
XX
KW   Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 27; SEQ ID No 34261; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a peptide/protein
CC   encoded by a single exon probe of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   133 AA;

  Query Match             11.7%;  Score 256;  DB 23;  Length 133;
  Best Local Similarity   68.0%;  Pred. No. 2.7e-13;
  Matches   51;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
          :||: |||| |||   || ||||||:|| : ||||||| || |||  ||||:||| |:|
Db     54 AESKNLKESVVPPTASIENKKQEREDKNWPILPPPVAETSVPPPSVAGIETPIQRILRSA 113

Qy    296 TIAGEPLGHCTFTIS 310
          ||||| | | | ||
Db    114 AIAGEPSGPCAFPIS 128


RESULT 7
ABG04611
ID   ABG04611 standard; Protein; 122 AA.
XX
AC   ABG04611;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #4602.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS68798.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 34970; 103pp; English.
```

```
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   122 AA;

  Query Match             11.6%;  Score 254;  DB 22;  Length 122;
  Best Local Similarity   59.1%;  Pred. No. 3.5e-13;
  Matches   52;  Conservative    7;  Mismatches   21;  Indels    8;  Gaps    1;

Qy     236 SESQILKESFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIPRTA 295
           :||: |||| |||| | || || |||:|| :|||||||| |  ||| ||||| |    :|
Db      33 AESKNLKESVVPPTAPTENKKQVREDKNWPIPPPPIAETSVLPPSVAEIETPKQTTLCSA 92

Qy     296 TIAGEPLGHCTFTIS--------PAFVH 315
           ||||||| ||| ||          |:|
Db      93 AIAGEPLGPCTFPISVRPDSNNPQQFIH 120


RESULT 8
ABB68342
ID   ABB68342 standard; Protein; 1178 AA.
XX
AC   ABB68342;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 31818.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL12445.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 31818; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1178 AA;

  Query Match             5.9%;  Score 130;  DB 22;  Length 1178;
  Best Local Similarity   19.8%;  Pred. No. 0.12;
  Matches   99;  Conservative   59;  Mismatches  180;  Indels  162;  Gaps   19;

Qy      49 QNPTLPSVSHRSPPGNAAVSVTGGDCHLPTEEEFGVLV-------QSMKCDTVRIKGVLQ 101
           | | :  |:   || : || :  :       |: | :        ||:   |  :
Db     416 QKPQVTCVTLVPPPQESPVSSSAAETLEKRPEQVGAISTSHESPRQSLSSPTDSVDSAKS 475

Qy     102 GPTTAPPLMTS--------------EGNVTAEDTEEAIRAFV---------------YAV 132
            |: :|   |               : |::|: |   : | |                 |
Db     476 TPSASPKPQTQVPMQPQLLQMRPHLQANLSAQ-TPVTVPAHVRVMRQQQAQIIPGQRLPV 534

Qy     133 AAASAAEAWHWRHLV-------LLSGQIHEPI-----------------------GSGG 161
           |::||| |   | :|        ::|:   |:                        | ||
Db     535 ASSSAAMATMTRSIVTSTAGTSTITGR---PVATTLNNSNPTVAQLQSMANAMNGAGGGG 591

Qy     162 NIINTNKG--------------------GRSC----QNPA-------------LPSPDQ 183
            :| |: |                    |:      | ||             |||
Db     592 QLIMTSSGQLLVIPTPSKQTTQHHRPGQPGQGVIIQQQQPAELHPQPGGGYIVSQPSPVA 651

Qy     184 SPSGNATTSVTRDN-------YHLLTEEEFGVWSQSMKWHSQNK----------SGGSVP 226
           : | ::::|    |        :| :   : |  :|:    | |:          :|| :
Db     652 AASSSSSSSTVILNSGGAKLLHHQIITSQAGQINQATSGGSGNQPQTVLLNTLPNGGYIV 711

Qy     227 VRGP-TQEPCSESQILKESFVPP------TTPKENNKQEREDENWRLPPPPVAE-----T 274
           : | || |    |||     ||       ::|    :  :   :    |||        :
Db     712 QQQPQTQSPQQAEQILAMPQPPPAQTLIISSPDTKRRARKRKSSVCHTPPPSGSPAKIIS 771

Qy     275 PVPSPSVTEIETPLQRIPRTATIAGEPLGHCTFTISPAFVHSVLNKRK------------ 322
           |  |||:     |      |  |  |     | :||     |:||
Db     772 PQISPSIPNQAPALLHQQAAAAAAAAPQFQQQFQLSPGIQGIVVNKPNPPQQPQTQQLLL 831

Qy     323 ------RQLELLLREVEWPGRGHMAATCCKLQVEGQDRT-MSLAAAPVREAPPPPTGASS 375
                :|: |: :::  |    |      ||::    | :     ||    | |    |
Db     832 QNGQILQQVNLIGQQLLMPAGLVMGPDATLLQIQNMPATSLMTPQGPVMLRTPSPQNKPS 891

Qy     376 --EPSVPA---LPGADPQRS 390
             ||      | ||: | |
Db     892 FISPSAGGQQYLVGANGQLS 911


RESULT 9
AAM38978
ID   AAM38978 standard; Protein; 641 AA.
XX
AC   AAM38978;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2123.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58134.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2123; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   641 AA;

  Query Match             5.9%;  Score 129.5;  DB 22;  Length 641;
  Best Local Similarity   21.9%;  Pred. No. 0.058;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |        |  |  :: ||   | | |  |||           | :
Db      90 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 140

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |            |  |: |   :      |  | |  :   :| |
Db     141 KSSGIPVKPV------------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 181

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
                 |  |    :   |||    |   | | |               ||: :: |:
Db     182 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS--------------LLSTKQGGLT 219

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV--------------PPTTPKENN 255
           :|  ::  || : |   |  :   | :  |   |               | :|||
Db     220 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 277

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE--------------TPLQRIP---RT 294
                 :   ||| |:   |:: |  ||:  ::               |   ::|    |
Db     278 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 337

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:    |    |      |   :|
Db     338 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 396

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
               ||| :|     ||   || |  | :|     |  :   |: :|
Db     397 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 441


RESULT 10
AAM38977
ID   AAM38977 standard; Protein; 698 AA.
XX
AC   AAM38977;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2122.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58133.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2122; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
```

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   698 AA;

  Query Match             5.9%;  Score 129.5;  DB 22;  Length 698;
  Best Local Similarity   21.9%;  Pred. No. 0.065;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |        |  |  :: ||   | | |  |||           | :
Db      90 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 140

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |              |  |: |   :      |  | |  :   :| |
Db     141 KSSGIPVKPV-------------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 181

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
               |  |    :   |||    |  | | |                ||: :: |:
Db     182 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS-------------LLSTKQGGLT 219

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV-------------PPTTPKENN 255
             :|  ::  || : |   |  :   | : |   |              | :|||
Db     220 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 277

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE--------------TPLQRIP---RT 294
               :   ||| |:   |:: |  ||:  ::               |   ::|    |
Db     278 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 337

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:     |    |        |  :|
Db     338 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 396

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
             ||| :|     ||   || |  | :|     |  :   |: :|
Db     397 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 441


RESULT 11
AAM39032
ID   AAM39032 standard; Protein; 698 AA.
XX
AC   AAM39032;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2177.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58188.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2177; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   698 AA;

  Query Match             5.9%;  Score 129.5;  DB 22;  Length 698;
  Best Local Similarity   21.9%;  Pred. No. 0.065;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |        |  |  :: ||   | | |  |||           | :
Db      90 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 140

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |              |  |: |   :      |  | |  :   :| |
Db     141 KSSGIPVKPV-------------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 181

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
               |  |    :   |||    |  | | |                ||: :: |:
Db     182 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS-------------LLSTKQGGLT 219

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV-------------PPTTPKENN 255
             :|  ::  || : |   |  :   | : |   |              | :|||
Db     220 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 277

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE--------------TPLQRIP---RT 294
               :   ||| |:   |:: |  ||:  ::               |   ::|    |
Db     278 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 337

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:     |    |        |  :|
Db     338 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 396

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
             ||| :|     ||   || |  | :|     |  :   |: :|
Db     397 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 441


RESULT 12
AAB94135
ID   AAB94135 standard; Protein; 705 AA.
XX
AC   AAB94135;
XX
DT   26-JUN-2001  (first entry)
XX
```

```
DE   Human protein sequence SEQ ID NO:14398.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama-T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 14398; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   705 AA;

  Query Match             5.9%;  Score 129.5;  DB 22;  Length 705;
  Best Local Similarity   21.9%;  Pred. No. 0.066;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |        |  |  :: ||   | | |  |||            | :
Db      90 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 140

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |              |  |: |   :       |  | |  :   :| |
Db     141 KSSGIPVKPV-------------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 181

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
               |  |    :   |||    |   | | |              ||: :: |:
Db     182 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS-------------LLSTKQGGLT 219

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV----------PPTTPKENN 255
              :|  ::  || : |   |  :   | :  |   |              | :|||
Db     220 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 277

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE----------TPLQRIP---RT 294
                 :  ||| |:   |:: |  ||:  ::              |   ::|    |
Db     278 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 337

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:    |    |     |   :|
Db     338 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 396

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
           ||| :|     ||   || |  | :|     |  :   |: :|
Db     397 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 441


RESULT 13
AAM38979
ID   AAM38979 standard; Protein; 706 AA.
XX
AC   AAM38979;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2124.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58135.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2124; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
```

```
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   706 AA;

  Query Match             5.9%;  Score 129.5;  DB 22;  Length 706;
  Best Local Similarity   21.9%;  Pred. No. 0.066;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |         |  |  :: ||   | | |  |||           | :
Db      90 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 140

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |               |  |: |   :       |  | |  :   :| |
Db     141 KSSGIPVKPV----------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 181

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
               |  |     :    |||    |   | | |                ||: :: |:
Db     182 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS--------------LLSTKQGGLT 219

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV--------------PPTTPKENN 255
             :|  ::  || : |   |  :   | :  |   |               | :|||
Db     220 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 277

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE--------------TPLQRIP---RT 294
                :  ||| |:   |:: |  ||:  ::                |   ::|    |
Db     278 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 337

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:     |    |       |  :|
Db     338 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 396

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
           |||  :|     ||   || |  | :|    |  :   |: :|
Db     397 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 441


RESULT 14
AAY85564
ID   AAY85564 standard; Protein; 1837 AA.
XX
AC   AAY85564;
XX
DT   07-JUL-2000  (first entry)
XX
DE   Human homologue of UNC-53 (Hs-UNC-53/1) sequence.
XX
KW   UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
KW   anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
KW   antisclerotic; antimetastatic; anti-arthritic; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          958..1014
FT                   /note= "this region is found to be absent when encoded by
FT                         a variant cDNA isolated from frontal cortex"
FT   Region          1033..1040
FT                   /note= "this region is found to be absent when encoded by
FT                         a variant cDNA isolated from frontal cortex"
FT   Region          1173..1175
FT                   /note= "this region is found to be absent when encoded by
FT                         a variant cDNA from Hela or colorectal
FT                         adenocarcinoma tissue"
FT   Misc-difference 1233
FT                   /label= Leu or Ser
XX
PN   WO9963080-A1.
XX
PD   09-DEC-1999.
XX
PF   02-JUN-1999;   99WO-EP03848.
XX
PR   03-JUN-1998;   98GB-0011962.
XX
PA   (JANC ) JANSSEN PHARM NV.
XX
PI   Luyten WHML,  De Raeymaeker MC,  Geysen JJGH,  Bogaert TAOE;
PI   Maerten LJS,  Verhasselt P,  Van De Craen M;
XX
DR   WPI; 2000-116370/10.
DR   N-PSDB; AAA07835.
XX
PT   Novel proteins and nucleic acids e.g. for treating neurodegeneration  -
XX
PS   Claim 93; Fig 1b; 146pp; English.
XX
CC   The invention provides vertebrate (human) protein homologue of a UNC-53
CC   protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or
CC   their plus ends. The UNC-53 sequences are used to promote neural
CC   regeneration, revascularization and wound healing; also for treating
CC   neurodegenerative disease, acute traumatic injury, fibrotic disease and
CC   autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
CC   polynucleotides can be used for recombinant production of the proteins,
CC   as a source of probes for detecting allelic variants and polymorphisms,
CC   for sequencing genomic DNA and for detecting UNC-53 expression; and as
CC   source of therapeutic antisense sequences. Cells that express the
CC   protein are used to identify regulators of cell shape, growth, motility
CC   and migration. They can also be used to identify proteins that are
CC   involved in signal transduction pathways also involving UNC-53, and to
CC   identify compounds that alter attachment of UNC-53 to microtubules. A
CC   target gene coupled to a UNC-53 encoding sequence may be used to deliver
CC   the target gene to a cellular microtubule or its plus ends. The present
CC   sequence represents the amino acid sequence of the first human homologue
CC   of UNC-53, designated hs-UNC-53/1.
XX
SQ   Sequence   1837 AA;

  Query Match             5.9%;  Score 129.5;  DB 21;  Length 1837;
  Best Local Similarity   21.9%;  Pred. No. 0.24;
  Matches   89;  Conservative   45;  Mismatches  165;  Indels  107;  Gaps   17;

Qy      41 GAQNLMTCQNP--------TLPSVSHRSPP---GNAAVSVTGGDCHLPTEEEFGVLVQSM 89
           |   |   : |         |  |  :: ||   | | |  |||           | :
Db     541 GRDRLSDAKKPPSGIARPSTSGSFGYKKPPPATGTATVMQTGGS---------ATLSKIQ 591

Qy      90 KCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAFVYAVAAASAAEAWHWRHLVLL 149
           |   : :| |               |  |: |   :       |  | |  :   :| |
Db     592 KSSGIPVKPV----------NGRKTSLDVSNSAEPGFLAPGARSNIQ---YRSL--- 632

Qy     150 SGQIHEPIGSGGNIINTNKGGRSCQNPALPSPDQSPSGNATTSVTRDNYHLLTEEEFGVW 209
               |  |     :    |||    |   | | |                ||: :: |:
Db     633 ----PRPAKSSSMSVT---GGRGGPRPVSSSIDPS--------------LLSTKQGGLT 670

Qy     210 SQSMKWHSQNKSGGSVPVRGPTQEPCSESQILKESFV--------------PPTTPKENN 255
             :|  ::  || : |   |  :   | :  |   |               | :|||
Db     671 PSRLKEPTKVASGRTTP--APVNQTDREKEKAKAKAVALDSDNISLKSIGSPESTPKNQA 728

Qy     256 KQEREDENWRLPPPPV---AETPVPSPSVTEIE--------------TPLQRIP---RT 294
                :  ||| |:   |:: |  ||:  ::                |   ::|    |
Db     729 SHPTATKLAELPPTPLRATAKSFVKPPSLANLDKVNSNSLDLPSSSDTTHASKVPDLHAT 788

Qy     295 ATIAGEPLGHCTFTISPAFVHSVLNKRKRQLELLLREVEWPGRGHMAATCCKL--QVEGQ 352
           :: :| ||  | || ||| : :: :    |   |:     |    |       |  :|
Db     789 SSASGGPLPSC-FTPSPAPILNINSASFSQGLELMSGFSVPKETRMYPKLSGLHRSMESL 847

Qy     353 DRTMSLAAA-----PVREAPPPPTGASSEPSVPALPGADPQRSAEL 393
           |||  :|     ||   || |  | :|    |  :   |: :|
Db     848 QMPMSLPSAFPSSTPV-PTPPAPPAAPTEEETEELTWSGSPRAGQL 892
```

```
RESULT 15
ABB65774
ID   ABB65774 standard; Protein; 828 AA.
XX
AC   ABB65774;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 24114.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL09877.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 24114; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   828 AA;

  Query Match             5.9%;  Score 128.5;  DB 22;  Length 828;
  Best Local Similarity   22.3%;  Pred. No. 0.099;
  Matches  100;  Conservative   44;  Mismatches  134;  Indels  171;  Gaps   22;

Qy      14 KQLGQALQASV-----SLSIITENQGKRCPFCGAQNLMTCQNPTLPSVSHRSPPGNAAVS 68
           :|  |  |||      |  |||  |         ||||    :|| |: |:
Db     488 QQQQQQEQASASGQAYSSQIITVNNLVGSYATAAQNL----SPTSPNESN---------- 533

Qy      69 VTGGDCHLPTEEEFGVLVQSMKCDTVRIKGVLQGPTTAPPLMTSEGNVTAEDTEEAIRAF 128
                       :|||:           ||||                 |:  : |
Db     534 ----------------MVQSV---------YSQGPT--------------PTQSPVHA- 552

Qy     129 VYAVAAASAAEAWHWRHLVLLSGQIHEPIGSGGNIINTNKGGRSCQNPALP-SPDQSPSG 187
           |  ||||                  |: ||    | |    |  :      :| |
Db     553 --GVGGASAAG-----------------GAAGNASAGNGGAPGAANQVVKRKRSVNPQG 592

Qy     188 NATTSVTRDNYHLLTE-------------EEFGVWSQSMKWHSQNKSGGSVPVRGPTQEP 234
           :       :|:    |              :|| :::: |    | |       |   |
Db     593 D-------ENFIRALEAVRTGGIGFCKAARLYGVNNRTL-WLEYKKRG------YPVSRP 638

Qy     235 CSESQILKE--SFVPPTTPKENNKQEREDENWRLPPPPVAETPVPSPSVTEIETPLQRIP 292
             :::::|:  :  |  ||  |   :  :|   :  || |||| ||   |   | |
Db     639 SIKARVVKQEPNLSPSPTPSTNQGDDNTNETLGMQIPPQAETPTPSLMCTSHHTGL---- 694

Qy     293 RTATIAGE-PLG---HCTFTI----------SPAFVHSVLNKRKRQLELLLREVEWPGRG 338
             : ||  | |   |    :          ||  |||:  :         | :|  | |
Db     695 --GSGAGSLPAGGNSHPAIGVMSLFDPRYMDSPGNVHSMTRQ---------RYIEATGGG 743

Qy     339 HMAATCC-------KLQVEGQDRTMSLAAAPVREAPPPPTGASSEP----SVPALPGADP 387
           | |          :  :    |:  ||| :          |:|||    |: |:| : |
Db     744 AGAGTGAGTGTINNTISSQAGTATLLQAAAVM---------AASEPAESLSISAMPVSVP 794

Qy     388 QRSAELLLL-----------AVTREGLE 404
             |   |:            ||  ||:
Db     795 MGMAHSFLINNFVTVAAPPVATVTSTGLQ 823


Search completed: December 25, 2002, 01:10:35
Job time : 72 secs
```